INFORMATION COLLECTION SYSTEM AND METHOD

Abstract

A user registers interests in order to obtain data relative to those interests. A plurality of web sites are searched based on the user's interests. An extraction rule mechanism is applied to HTML documents acquired. A vocabulary information processing mechanism reads an ontology based on information received to obtain vocabulary information. An inference processing mechanism executes an inference operation based on axiom rules. An extracting position information identifying section extracts data objects relying on HTML document tags with respect to the acquired HTML documents, based on the extraction rules, the vocabulary information processing mechanism, and the inference processing mechanism.